Agricultural and Forest Meteorology 184 (2014) 82–97

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

journal homepage: www.elsevier.com/locate/agrformet

On seeing the wood from the leaves and the role of voxel size in determining leaf area distribution of forests with terrestrial LiDAR

Martin Béland [a,b,*], Dennis D. Baldocchi [a], Jean-Luc Widlowski [c], Richard A. Fournier [b], Michel M. Verstraete [c,1]

[a] Department of Environmental Science, Policy and Management, University of California, Berkeley, Berkeley, CA 94720, United States
[b] Centre d'Applications et de Recherches en Télédétection (CARTEL), Université de Sherbrooke, Sherbrooke, Québec, Canada J1K 2R1
[c] European Commission-DG Joint Research Centre, Institute for Environment and Sustainability, Ispra, VA I-21027, Italy

ARTICLE INFO

Article history:
Received 29 June 2013
Received in revised form
10 September 2013
Accepted 14 September 2013

Keywords:
Terrestrial LiDAR
Voxel
3-D leaf area distribution
Leaf area density
Leaf area index
Forest canopy structure

ABSTRACT

Terrestrial LiDAR scanners have been shown to hold great potential for estimating and mapping three dimensional (3-D) leaf area distribution in forested environments. This is made possible by the capacity of LiDAR scanners to record the 3-D position of every laser pulse intercepted by plant material. The laser pulses emitted by a LiDAR scanner can be regarded as light probes whose transmission and interception may be used to derive leaf area density at different spatial scales using the Beer–Lambert law or Warren Wilson's contact frequency method among others. Segmenting the canopy into cubic volumes –or voxels– provides a convenient means to compute light transmission statistics and describe the spatial distribution of foliage area in tree crowns. In this paper, we investigate the optimal voxel dimensions for estimating the spatial distribution of within crown leaf area density. We analyzed LiDAR measurements from two field sites, located in Mali and in California, with trees having different leaf sizes during periods with and without leaves.

We found that there is a range of voxel sizes, which satisfy three important conditions. The first condition is related to clumping and requires voxels small enough to exclude large gaps between crowns and branches. The second condition requires a voxel size large enough for the conditions postulated by the Poisson law to be valid, i.e., a turbid medium with randomly positioned leaves. And, the third condition relates to the appropriate voxel size to pinpoint the location of those volumes within the canopy which were insufficiently sampled by the LiDAR instrument to derive reliable statistics (occlusion effects). Here, we show that these requirements are a function of leaf size, branching structure, and the predominance of occlusion effects. The results presented provide guiding principles for using voxel volumes in the retrieval of leaf area distributions from terrestrial LiDAR measurements.

1. Introduction

Tree foliage properties are critical to describe the interactions between the land surface and the atmosphere, in particular the rates of radiation absorption, precipitation interception, and photosynthetic activity in individual plants and whole canopies (Baldocchi and Harley, 1995; Parker, 1995). As such, the leaf surface generated by trees is of great interest to the global change research community because of their role in the terrestrial part of the global carbon cycle (Sellers et al., 1997; Asner et al., 2003). Moreover, studies investigating the implications of light transmission and interception for wood production, species competition, ecosystem and agro-ecosystem dynamics, as well as biodiversity, rely on descriptions of the spatial distribution of leaves (Norman and Welles, 1983; Cannell, 1989; Wang and Jarvis, 1990; Parker, 1995; Medlyn, 2004; Asner et al., 2008; Pretzsch, 2009).

In this contribution, leaf area refers to half of the total foliage surface area (only one side of each leaf is considered). This concept only applies to leaves that can be flattened, and may be less effective or more complicated to use with needles and leaves of arbitrary shapes and forms. The leaf area density [m^2/m^3] is generally used to quantify canopy foliage within a given volume. Leaf area density

* Corresponding author at: Department of Environmental Science, Policy and Management, University of California, Berkeley, 130 Mulford Hall, Berkeley, CA 94720, United States. Tel.: +1 510 642 2421; fax: +1 510 643 5438.
E-mail addresses: mbeland@berkeley.edu, martin.beland@usherbrooke.ca (M. Béland).
[1] Present address: South African National Space Agency (SANSA), Pretoria 0087, South Africa.

0168-1923 © 2013 The Authors. Published by Elsevier B.V. Open access under CC BY-NC-SA license.
http://dx.doi.org/10.1016/j.agrformet.2013.09.005

can, in theory, be specified at any given spatial scale (e.g., at the sub-crown level or at the plot level) and can be estimated *in situ* using different approaches that have been classified as 'direct', 'semi-direct', or 'indirect'. Direct methods involve the counting and measuring of leaves; this application is time consuming (Norman and Campbell, 1989). In contrast, semi-direct methods include the inclined point quadrat method of Warren Wilson (1960) that counts the number of contacts of leaves with probes inserted into the vegetation canopy. Numerous passive optical indirect techniques have been developed for leaf area density or leaf area index (LAI) estimation, including hemispherical photography (see Jonckheere et al. (2004) for a review). These methods rely on the Beer–Lambert law of light transmission through a turbid medium which was adapted to canopies (Monsi and Saeki, 1953, 2005); these methods are limited in the spatial explicitness of their estimates and in their accuracy (Weiss et al., 2004). The latter is partially due to the need to correct for foliage clumping (Nilson, 1971; Chen and Black, 1991) as well as to remove the contribution from non-photosynthetic material (wood) (Kucharik et al., 1998). It may be useful to note that the results obtained from particular measurement protocols tend to yield different results when applied simultaneously to the same canopy, particularly in heterogeneous canopies (Ryu et al., 2010).

Since the bulk of the exchanges of radiation, heat, water vapour and biogeochemicals between the atmosphere and the vegetated land surface are mediated through leaves, the latter play an important role in controlling the local microclimate in the canopy. The fine-scale spatial organization of leaves determines the light regime within a canopy, and is an important consideration when modeling canopy functions because of the non-linear leaf-level photosynthetic and stomatal conductance response to light. Information on leaf area distribution is required to properly account for the transfer of radiation within canopies (Hutchison et al., 1986). Jarvis et al. (1985) showed that satisfactory upscaling of photosynthetic rates from individual tree crowns to the canopy stand requires being able to simulate the radiative regime within individual tree crowns. Great progress has been made in modeling light transfer in discontinuous canopies since the early works of Jackson and Palmer (1972, 1979). Current 3-D radiative transfer models used to compute mass and energy transfer – e.g. MAESTRO/MAESTRA (Norman and Welles, 1983; Wang and Jarvis, 1990; Medlyn, 2004), DART (Gastellu-Etchegorry et al., 1996), FOREST (Cescatti, 1997), RATP (Sinoquet et al., 2001), RAYTRAN/RAYSPREAD (Widlowski et al., 2006), CANOAK-FLiES (Kobayashi et al., 2012) – are capable of utilizing spatially explicit descriptions of leaf area distribution in their computations. However, several studies have highlighted the remarkable efforts required to characterize the distribution of leaf area within actual tree crowns with existing field methods (Fukai and Loomis, 1976; Koike, 1985; Cohen and Fuchs, 1987; Wang et al., 1990; Whitehead et al., 1990; Landry et al., 1997; Sinoquet and Rivet, 1997; Pearcy and Yang, 1998; Sinoquet et al., 1998; Iio et al., 2011). Hence, making use of three-dimensional radiative transfer models has been impractical for routine application.

LiDAR scanners have recently emerged as a promising tool for deriving vegetation structure descriptions at various levels of detail. LiDAR scanners use a large number of laser pulses emitted in the visible or near-infrared part of the spectrum within the instrument's field of view. When a pulse comes into contact with an object, part of that energy is reflected back toward the instrument and triggers the recording of its distance (either using time of flight or phase displacement) and intensity. Knowing the direction of the emitted laser pulse, the latter can be used to position a point in 3-D space (a set of points is referred to as a point cloud). Discrete return LiDAR systems are either able to record the position of only one contact between the pulse and plant material – referred to as single discrete return systems–, or several contacts – referred to as multiple discrete returns systems. Full-waveform systems can record a highly detailed account of the reflected energy received as a function of time (at the cost of longer scanning times due to the increased volume of data to digitize and store). Different LiDAR systems have different pulse cross-sectional sizes that increase with distance from the instrument. The size of the cross-sectional laser pulse area is typically quantified using the mean diameter; the term 'diameter of the laser footprint' is also used. Note that there are a number of ways to measure a laser pulse diameter, the simplest way being the full width at half maximum. Current typical sizes are 0.01–0.2 m at 20 m for terrestrial LiDAR, 0.2–0.9 m for discrete return airborne LiDAR, 8–70 m for full-waveform airborne LiDAR, and 70–100 m for spaceborne LiDAR. Comparing these with the sizes of tree parts, one can see that the energy from a laser pulse can be only partly intercepted by plant parts, allowing a remaining fraction of energy to keep traveling deeper into the canopy and make contact with material located further along its path.

Terrestrial LiDAR measurements are generally made from an instrument placed on a survey tripod about 1.5 m above the ground. Their usage for estimating leaf area stems from a very high spatial resolution (typically about half a centimeter between consecutive laser pulses at 20 m from the instrument) and a relatively small laser footprint size with respect to the typical dimensions of leaves and other tree organs. The active nature of terrestrial LiDAR instruments provides two valuable pieces of information not readily available to passive methods like hemispherical photography: The first is a measure of range between the instrument and the material which intercepted the emitted light, and the second is a measure of the intensity of the light which was reflected by the material when the laser pulse hit it. The availability of these two pieces of information can help solve issues burdening the above mentioned indirect field methods of leaf area and LAI retrieval that are based on passive optical remote sensing. More specifically, the range – and 3-D positions – can be utilized to segment the canopy into small volumes (e.g., voxels) where clumping is negligible such that the Beer–Lambert law can be used at these scales without having to correct for clumping effects (Nilson, 1971; Lang and Xiang, 1986). In addition, the intensity of the reflected LiDAR pulse can be used to distinguish between return signals that were the result of an interaction with wood or with foliage material, e.g., Béland et al. (2011). The later aspect is, however, only possible if there is a significant difference in reflectivity between the wood and foliage material at the wavelength of the terrestrial LiDAR instrument. This type of approach for estimating leaf area density can be classified as an active optical indirect method.

So far, only a handful of studies have investigated the use of terrestrial LiDAR scanners for estimating tree and canopy level leaf area density compared to those focusing on the retrieval of leaf area density by other means (Lovell et al., 2003; Hosoi and Omasa, 2006, 2007; Moorthy et al., 2008; Takeda et al., 2008; Hosoi and Omasa, 2009; Jupp et al., 2009; Van der Zande et al., 2009; Hosoi et al., 2010). Among these efforts, Béland et al. (2011) showed that terrestrial LiDAR measurements open new possibilities for studying the 3-D foliage distribution of large trees and derived spatially explicit estimates of leaf area density. Deriving spatially explicit measures of leaf area, instead of spatially representative leaf area estimates, at the crown and stand level is especially advantageous in heterogeneous environments, like savannas, where adequate spatial sampling can be challenging (Ryu et al., 2010). It also provides much richer information about canopy structure, such as vertical and horizontal organization of plant material. When deriving leaf area within the confines of geometric volumes like voxels, the size of the volume needs to be chosen carefully as it can significantly influence the leaf area estimates. To the best of our knowledge, there are no clear guidelines currently available for selecting an appropriate voxel size. In a previous paper (Béland et al., 2011), it was noted that when using voxels larger than the diameter of the laser pulse

cross section, the voxel size should "provide a balance between a large enough voxel size to satisfy the assumptions of random distribution of small scatterers, and small enough to correctly identify the occluded areas and account for foliage clumping". However, the factors influencing this balance need further attention.

Other typical issues that have to be dealt with when extracting leaf area information from terrestrial LiDAR data are: (1) accounting for displacement of branches and foliage due to the presence of wind during the scans, (2) distinguishing between return signals from wood and foliage, (3) accounting for the underestimation of gap fraction due to the finite cross section of the laser pulses, and (4) avoiding systematic biases from canopy volumes not sampled by the laser scans due to occlusion by other canopy elements. The term occlusion here refers to the fact that, unlike the needles used in the point quadrat method of Warren Wilson, the laser pulses are blocked – at least partially – by the leaves, keeping the pulses from making contact with leaves located further along their path. If occlusion reduces the number of pulses traveling into a given cubic volume (known as a voxel) too much, then few meaningful statistics may be derived from these data. Using a voxel size significantly larger than the LiDAR laser pulse cross section is a convenient approach to dealing with occlusion effects. The pulses entering the voxel volume can be interpreted as an ensemble of probes from which statistics on their interception by leaves can be derived. This approach provides a means to identify voxels sampled by too few pulses to provide a reliable leaf area estimate. Béland et al. (2011) found that a minimum of 15% of the voxel volume explored by the laser pulses was required to yield reliable leaf area density values.

An important advantage of using terrestrial LiDAR measurements for deriving leaf area density lies in the potential for discriminating between laser pulses that were intercepted by wood and foliage. Successful distinction of laser pulse returns originating from wood and foliage material is highly dependent on an understanding of the interactions between laser pulses – driven by the LiDAR instrument's electronics and the canopy optical properties – and the canopy material. Béland et al. (2011) have shown that using scans acquired in leaf-on and leaf-off conditions can help understanding these interactions, and that the intensity of the returns can be used to make the distinction. However, they also pointed out that the intensity of a laser return is a function of three variables: (1) the range to the target, (2) the spectro-directional properties of the target and its orientation, and (3) the portion of the laser pulse hitting the target. The first can be corrected for by normalizing the intensity so that an apparent reflectance independent of range is derived. The second can be estimated using a field spectrometer (the discrimination of foliage from wood is based on a reflectance difference between the two at the wavelength of the LiDAR instrument). The third is a potential source of errors, as the partial illumination of a bright target can yield the same measurement as a more complete illumination of a darker target. When a partial contact between a pulse's energy and a target occurs, part of the energy keeps traveling forward without disturbance, but that proportion is unknown, creating an uncertainty about the target's reflectance even if Lambertian reflectance is assumed. This results from two processes: (1) the radiation that is not returned to the instrument is not measured, and hence it is not possible to know whether it has been absorbed or scattered (in particular transmitted), and (2) independently, the reflectance that is actually measured depends on the anisotropy of the surface actually illuminated. Note that both these points apply whether the target is fully or partially illuminated by the pulse. Recently available multiple discrete returns or full-waveform terrestrial LiDAR instruments could provide information useful toward reducing this uncertainty.

In this study, we first test the hypothesis that measurements from a multiple return terrestrial LiDAR scanner instrument can reduce uncertainty related to the classification of laser pulse returns originating from hits on wood and foliage material. This is necessary to get better statistics about the presence of foliage in the voxels. Then we focus on the main goal of the study, which is to investigate an appropriate voxel volume size for deriving spatially explicit leaf area estimates at the voxel scale. We expect that this size is related to the dimensions of the tree leaves, and to the spatial arrangement of leaves around branches. This assumption is based on the Beer–Lambert law which requires that volumes are filled with a large number of randomly distributed small scatterers. If leaves are relatively large compared to the sampling volume, or if the volume is large and includes empty spaces between plant parts, then the leaf area estimates deviate from the true value, as was shown by Lang and Xiang (1986). Here we use terrestrial LiDAR measurements taken at two different savanna sites with trees having different leaf sizes to investigate the optimal voxel size for deriving leaf area estimates.

2. Material and methods

2.1. Sites characteristics and LiDAR measurements

Two study areas were used for this research. The first is located near Segou in Mali. Here terrestrial LiDAR measurements were acquired in both leaf-on and leaf-off conditions and direct measurements of the leaf area are available for validation purposes. The second area is located in California, USA. Here LiDAR measurements in tree leaf-on and leaf-off conditions using a multiple-return instrument are available. They were used to explore the feasibility of better distinguishing between wood and foliage laser returns. In addition, the LiDAR measurements from both sites were used to assess the links between leaf size and branching structure, and to estimate appropriate voxel sizes to be used for computing leaf area estimates.

The Malian study area lies within a broad-leaved savanna of central Mali dominated by shea trees (*Vitellaria Paradoxa*). The area is generally flat, located at latitude 13.293° North and longitude 6.551° West, at an altitude of 290 m. The mean annual rainfall is in the 600–1000 mm range. For this study, a set of six shea trees were selected within an area of 50 m × 50 m to carry out the LiDAR measurements as well as direct measurements of leaf area. Within an area of 250 ha surrounding the study area, 25 zones of 1 ha were randomly selected to compute an average tree density of 17 trees ha^{-1}. For 25 randomly selected trees located close to the six laser scanned trees, the mean tree height was 8 m (standard deviation, (S.D.): 1.3 m). The shea tree leaves are elliptical in shape and relatively flat, with some slight undulation at the edges (see Fig. 1). Selected structural properties of the trees are presented in Table 1 and an illustration of LiDAR measurements from one of the tree is shown in Fig. 2. The leaf area values for the six trees were obtained by harvesting the tree leaves (the process required 20 man-days) and using a fresh-weight-to-area-ratio method (Norman and Campbell, 1989). Béland et al. (2011) provide further details on the observation protocol and measurement accuracy.

LiDAR measurements at the Mali site were acquired in February 2009 with an ILRIS-3D (Optech inc., Ontario, Canada) instrument in first return mode in both leaf-on and leaf-off conditions (after leaf harvesting). The measurements took place on days with very low to imperceptible wind conditions. The ILRIS-3D terrestrial LiDAR instrument emits laser pulses at a wavelength of 1535 nm within a maximum window size of 40° × 40° field of view and uses the pulse's travel time to compute the distance to targets. The minimum angle between consecutive laser pulses emitted by the ILRIS-3D is 26 μrad horizontally as well as vertically. The diameter of the laser pulse is 12 mm when leaving the instrument and

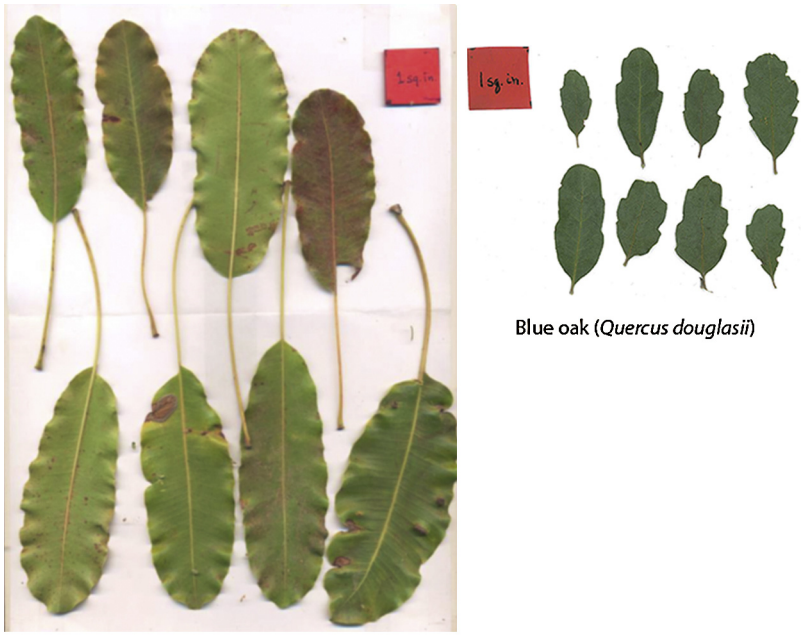

Fig. 1. Digitized images of shea tree leaves (left) and blue oak leaves (right) illustrating the difference in leaf size between the two species. The red rectangles provide an appreciation of scale, each being one square inch in area.

Table 1
Shea tree properties.

Tree #	DBH[a] [cm]	Height [m]	Leaf area [m^2]	Mean leaf size	
				Width [cm]	Length [cm]
1	31	7.6	29	4.1	11.8
2	34	7.9	105	3.8	8.7
3	39	7.8	104	4.4	10.8
4	34	7.2	101	3.6	10.1
5	36	7.3	126	3.9	10.2
6	41	8.7	530	4.0	12.7

[a] Trunk diameter at breast height.

increases slightly with distance (pulse divergence is 0.17 mrad, resulting in a footprint of about 15 mm at a distance of 20 m from the instrument). The pulse's energy is not uniformly distributed within its cross section, but follows a Gaussian shape (Optech inc., personal communication, 2009). Two LiDAR scans from opposite directions were taken for each shea tree with the instrument setup on a survey tripod about 1.7 m above ground and at a distance of approximately 20 m from the trunk. The angular resolution selected for these scans was approximately 260 μrad between consecutive laser pulses, yielding a distance between laser pulses of about 5 mm at a distance of 20 m from the instrument. Using this configuration, each scan took 20–25 min to complete.

The Californian study area is located near Ione, CA, on the Tonzi ranch, which is part of the AmeriFlux and FLUXNET micrometeorological observations networks (38.43° N, 120.96° W, elevation: 177 m). The area is generally flat and receives on average about 560 mm of rain annually, mostly during the winter months (Chen et al., 2006). The site is largely dominated by blue oaks (*Quercus douglasii*), with a density of about 144 trees/ha and a mean tree height of about 9.5 m (Ma et al., 2007). The blue oak leaves are elliptical in shape with small lobes (see Fig. 1). From a sample of 335 leaves collected from different trees at heights of 2 and 5 m above the ground, the average leaf size was about 3.5 cm (S.D. = 0.9 cm) along the midrib and 2 cm (S.D. = 0.7 cm) across. Ten isolated individual

Fig. 2. Illustrations of the LiDAR point clouds for shea tree (left) and blue oak tree (right). In each case, the coloring of foliage and wood parts is done using the apparent reflectance (normalized for distance) derived from the intensity of the reflected laser energy.

trees in addition to a 20 m × 20 m area containing a cluster of trees were scanned using a Riegl VZ-400 (Riegl inc., Austria) in March 2012 (leaf-off condition) and in May 2012 (leaf-on condition). An illustration of the measurements from one of the trees is shown in Fig. 2. The number of trees selected for scanning was based on the available time to complete the field measurements. The selection of the ten trees was driven by the wish to obtain a diversity of crown sizes and shapes. The average wind speeds were 0.9 m/s (0.7 m/s) 2 m above ground and 2.1 m/s (1.9 m/s) 23 m above ground for the March (May) measurement days (ORNL DAAC, 2012). At these wind speeds, leaf motion was barely visible, and was considered to have little effect on the measurements.

The VZ-400 terrestrial LiDAR instrument used at Tonzi emits laser pulses at a wavelength of 1550 nm and can cover 360° azimuthally and from −40° to +60° in elevation angle. The instrument has three important features which are not common among other terrestrial LiDAR scanners: (1) it can perform scans rapidly at a speed of 122,000 measurements per second, (2) it offers the possibility of using specially designed targets to make multiple scan co-registration significantly more efficient when combined with the RiScan Pro software (Riegl inc., Austria), and (3) the instrument can record full-waveform data (i.e., all pulse echoes are recorded at a small temporal interval) but this comes at a cost in acquisition time. The diameter of the laser pulse when leaving the instrument is 7 mm, and the pulse divergence is 0.3 mrad. The trees were scanned from a distance of about 10–12 m. At this distance, the pulse cross section diameter for the Riegl VZ-400 is about 1 cm. The overall plot area was scanned from six different locations. All leaf-off and leaf-on scans were carried out from the same positions using the same angular resolution of about 280 μrad. Depending on the coverage area settings used, each scan took between 2 and 6 min to complete.

2.2. Distinguishing noise, wood, and foliage from terrestrial LiDAR measurements

It is essential to screen out pulse returns from woody elements to estimate the leaf area density accurately. Otherwise the wood returns will artificially increase the apparent foliage content in the volume sampled by the laser pulses (thus providing estimates of plant area density instead of leaf area density). The LiDAR measurement files used here consist of a list of point coordinates in 3-D space each tagged with an identifier describing the nature of the point, i.e., leaf, wood, or noise, as described below. Noise points typically occur in two situations: (1) when two or more targets spatially close together are partially hit by the laser pulse, producing a point located somewhere in between the targets (this is commonly called an 'air point' and results from limitations in the instrument's electronics to distinguish the signal from the two targets) and (2) when the LiDAR instrument is able to measure very small quantities of reflected energy due to very small portions of a pulse hitting a target at close range (this typically occurs when very sensitive instruments – designed to measure targets at very large distances – are used to scan nearby objects). The Riegl VZ-400 records, for each return, the deviation of the Gaussian shaped signal received, providing an indication of the probability that the return is an 'air point'. The RiScan Pro processing software was used to filter out points with a high deviation. These points are considered as 'noise' in the input files for processing.

Distinguishing leaf from wood targets using the intensity of the reflected pulse requires significant differences in the optical properties of these components at the wavelength used by the LiDAR instrument. In this study, the terrestrial LiDAR instruments both operated in the 1535–1550 nm range. This wavelength range is close to a water absorption window such that foliage reflects less energy than wood. In fact, measurements made with an ASD Fieldspec Pro spectrometer (ASD, Boulder, CO, USA) at the California sites indicated a difference of about 0.15–0.20 in reflectivity (the average leaf reflectance at 1550 nm was computed as 0.30 (S.D.: 0.027) on the basis of 18 samples collected from six trees at heights of 2 and 5 m above ground). The wood-leaf separation approach is based on identifying an appropriate intensity threshold value. More specifically, it requires that the intensity of the return signal be normalized so that it becomes independent of the distance between the instrument and the target, e.g., Béland et al. (2011) and Pfennigbauer and Ullrich (2010). These methods (so far) have assumed Lambertian scattering properties for the targets and also a complete illumination of the target by the laser beam. As such the normalized reflectance delivered by these methods must be termed an "apparent" reflectance.

In this framework, for example – if we assume Lambertian reflectance – 55% of a pulse cross section hitting a branch having a reflectance of 45% will generate a return with an apparent reflectance of about 25%; the same value would be generated from an entire pulse hitting a leaf having a reflectance of 25%. In this example, the two returns cannot be distinguished on the basis of apparent reflectance, and this problem is more likely to occur where a large part of the small branches and twigs are visible to the LiDAR instrument – e.g. if tree leaves are small and/or the canopy LAI/PAI ratio is low (LAI is an indicator of the amount of foliage per ground area, and PAI (plant area index) is an indicator of the area of wood and foliage per ground area). However, there is one additional difference between the two cases from the example above which could help reducing the uncertainty: in the first case, 45% of the laser pulse energy keeps moving forward and is able to make contact with other targets located along the path of the pulse, whereas in the second case only the energy transmitted through the leaf keeps moving forward (leaf transmittance typically being in the same order of magnitude as reflectance (Jacquemoud and Ustin, 2001)), and reflected energy from contact with other target further along the path of travel would need to be transmitted a second time through the same leaf, which renders low the number of such occurrences given typical LiDAR instrument sensitivity.

The Riegl VZ-400 instrument is able to record multiple contacts that a laser pulse may experience as it travels through the canopy. This offers new perspectives for improving the identification of those data points representing leaves. To assess the usefulness of this new capability, we analyzed the apparent reflectance histograms from scans of ten blue oak trees acquired in leaf-on and leaf-off conditions.

Fig. 3 shows typical histograms from a blue oak tree crown scanned using the Riegl VZ-400. In gray are the data points relating to the number of pulses for which one or multiple returns were recorded ('multiple and single'), and in black are the data points relating only to those pulses for which one single return was generated ('single only') – i.e. ignoring those pulses that made contact with more than one target. It can be seen that pulse returns having lower reflectance are more likely to make contact with multiple targets. The two peaks correspond to returns that originate predominantly from leaves (centered at ∼21% apparent reflectance) and bark (centered at ∼47% apparent reflectance). The large amount of very low apparent reflectance returns in the multiple and single data points can be explained by the fact that LAI/PAI ratios are relatively low for the blue oak such that a large amount of laser pulses graze the multitude of small branches and twigs which causes small amounts of energy to be scattered back to the LiDAR instrument. One will notice in Fig. 3 the average apparent reflectance of the peak representing the leaves is (at ∼21%) smaller than the mean reflectance measured with the ASD instruments (∼30%). This can be due to two partial hits and non-Lambertian scattering of the foliage. For example, if one assumes that the leaves are Lambertian, and that multiple scattering is negligible then on average a pulse

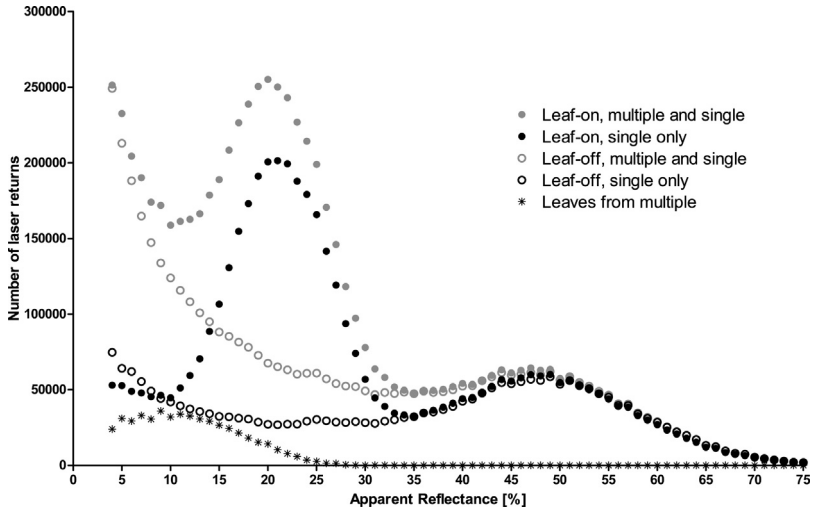

Fig. 3. Histograms of the TLS point clouds corresponding to leaf-on (filled disk) and leaf-off conditions (open disk), and to laser pulses making contact with one single target (black disk) vs pulses making contact with multiple targets along their paths (gray disk). Also shown is the difference between the number of pulses making contact with multiple targets in leaf-on and leaf-off conditions (star symbol), corresponding roughly to the number of pulses hitting leaves. An apparent reflectance of 4% was the lowest detectable value in the present case.

must have hit the blue oak leaf with only about 70% of its energy (21%/30%). On the other hand it is known that leaves scatter light anisotropically. Balduzzi et al. (2011) applied a FARO instrument operating at 785 nm on pear trees leaves and recorded a decrease in the intensity of the LiDAR return signal of ~10% as the angle between the laser pulse and the leaf normal was increased to about 80°. Depending on the specular component of foliage these results are, however, likely to be different.

The leaf-off condition histograms in Fig. 3 shows that a number of LiDAR returns, corresponding to partial hits on wood surfaces, fall within the 15–30% apparent reflectance range that nominally corresponds to return signals from leaves. These returns cannot be distinguished from the leaf returns on the sole basis of apparent reflectance. The occurrence of wood returns in the 15–30% apparent reflectance range can be reduced by excluding the pulses which made contact with more than one target, i.e. using the 'single only' returns. By doing so, however, one also discards a number of valid returns from leaves. We thus need to determine how many valid returns are discarded and adjust the reflectance threshold used to separate leaf from bark returns accordingly. For every value of apparent reflectance and for each tree, we computed the difference between the number of pulse returns that made contact with multiple targets in leaf-on and in leaf-off conditions. This provides an estimate of the amount of pulse returns from leaves that made contact with multiple targets (star symbols in Fig. 3). The difference between 'multiple and single' and 'single only' returns in leaf-off condition was also computed, providing an estimate of the proportion of pulse returns from wood contributing to the classification error which is removed by using 'single only' returns. This information will be used later on to assess the improvement in distinguishing wood from foliage returns.

2.3. Implications of laser pulse cross-sectional size

When using laser pulses emitted by a terrestrial LiDAR instrument as a proxy for light transmission through a medium, the size of the pulse's cross section should be considered carefully since partial interception of pulses can result in a significant underestimation of gap fraction (Danson et al., 2007). LiDAR instruments typically provide two values used to determine the size of the laser cross section as a function of distance from the instrument: the initial cross section size as the pulse leaves the instrument, and the beam divergence describing the increase in size with distance. These values can vary significantly between LiDAR instruments. For example, the two commercial systems used in this study had footprint diameters ranging from 13–15 mm at a distance of 20 m from the instrument; in contrast, other existing systems have significantly larger cross-sectional diameter upon exiting the instrument as well as higher pulse divergence, resulting in a footprint in the order of 100–300 mm in diameter at a distance of 20 m. If the laser pulse cross-sectional size is sufficiently small, its effect on leaf area density estimates should be accounted for. This is similar to Warren Wilson (1963) who showed that a correction was needed to account for the size of probes in the point quadrat method. If the laser pulse size is too large to allow such correction to be made, one must rely on an estimation of the gap fraction based on modeling of an apparent reflectance to which multiple targets contributed. Fig. 4 illustrates both of these situations. In the context of terrestrial LiDAR scans, Béland et al. (2011) provided a correction function ($H(\theta_L)$) that accounted for the pulse size (at a given distance), the sensitivity of the instrument's sensor, as well as the leaf size and shape.

When using the wood-leaf separation approach described in Section 2.2 above, i.e. where pulse returns are classified as noise on the basis of making contact with multiple targets, a correction function like $H(\theta_L)$ can no longer be used to correct for pulses making contact with a leaf while having their center outside the leaf edge. However, if one defines the spectrometer measured leaf reflectance as the nominal true value and furthermore imposes directional invariance of the retro-reflection intensity, then all laser returns with an apparent reflectance value of less than 50% of the nominal (spectrometer measured) leaf reflectance must have their center fall outside the area of the leaf. Hence, for leaves having a nominal reflectance of 30%, this framework suggests that all pulse returns having less than 15% apparent reflectance must have their pulse center outside the area of the leaf. For such cases, the pulse should thus not be considered as a hit on a leaf, but rather should be classified as 'noise' in the computation of the leaf area density estimates. Using the apparent reflectance to correct for the pulse footprint size has the additional advantage of taking into account the pulse divergence, which allows correcting for the effect of pulse cross-sectional size hitting targets at any distance from the instrument.

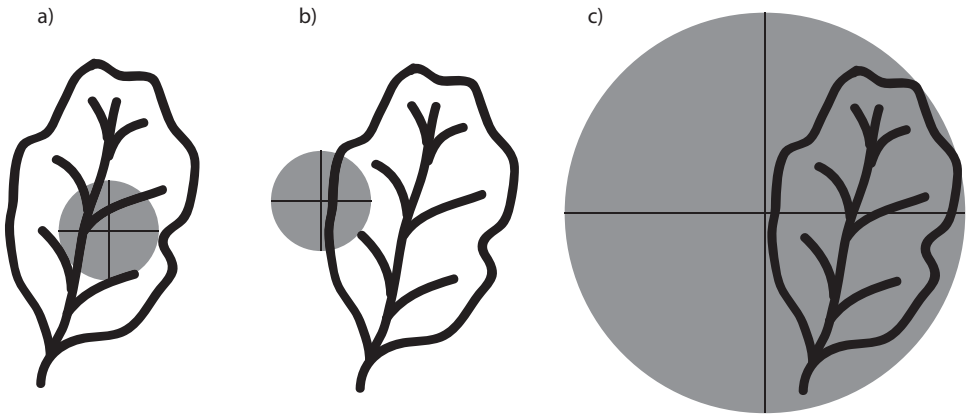

Fig. 4. Illustration showing the laser pulse footprint size (in gray) vs the leaf size. (a) The pulse footprint is smaller than the leaf and its center falls within the leaf area. (b) The pulse center falls outside the leaf area, in which case the pulse should be excluded from gap fraction computations. (c) The pulse footprint is significantly larger than the leaf, in which case using the intensity of the reflected energy to determine if the pulse center is located within the leaf area becomes problematic.

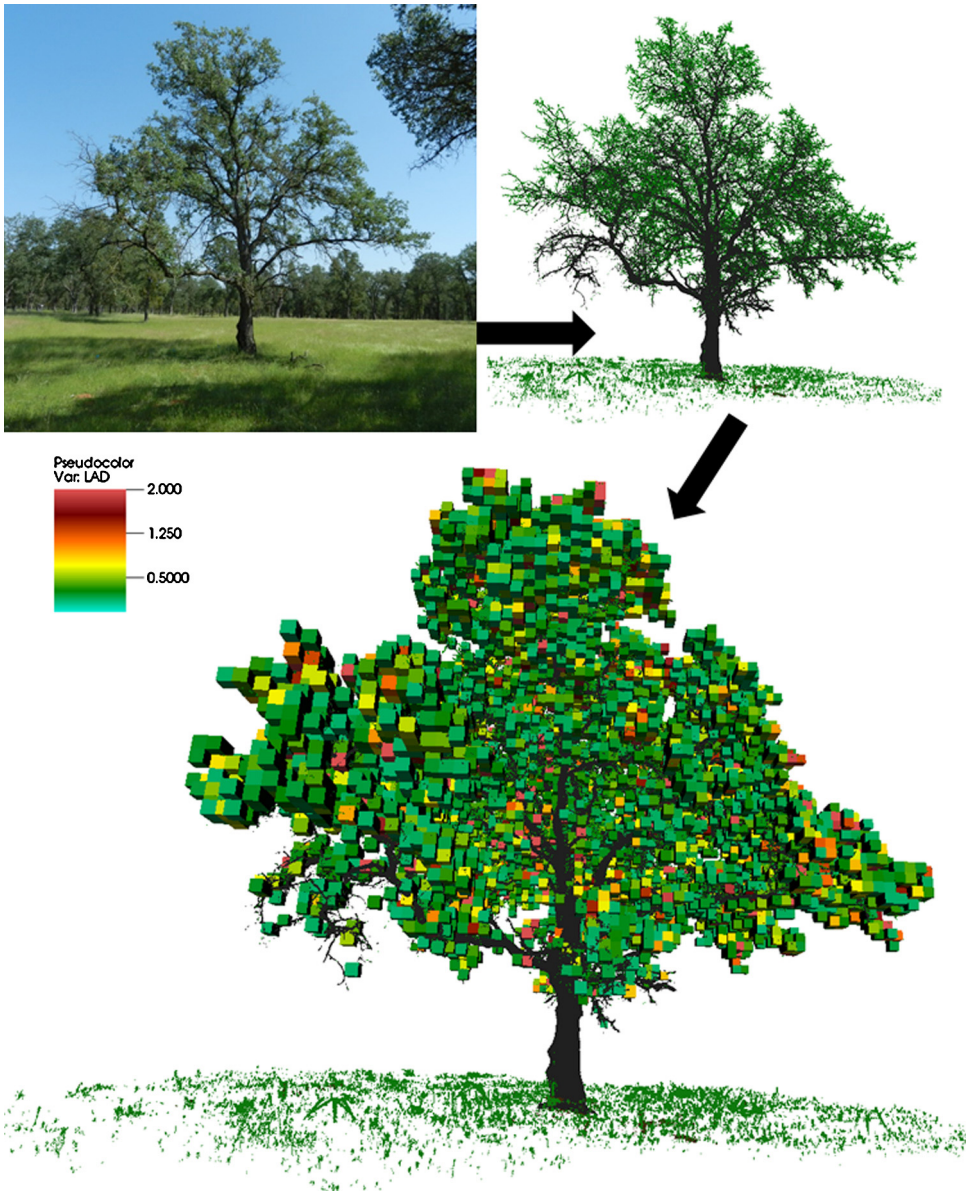

Fig. 5. Illustration of the concept of using voxel volumes for segmenting 3-D space and describing the distribution of leaf area within tree crowns. Statistics relating the interactions of the LiDAR pulses with plant material within each voxel are used to derive the density of leaf material therein. In this example, nearly 20,000 voxel of 15 cm in side length were used to describe the distribution of leaf area. The apparent difference in voxel sizes in the image results from differences in distance from the point of view location.

2.4. Implications of voxel size for leaf area estimates

Our approach for analysing the LiDAR data uses cubic volumes (voxels) as the basic element for deriving leaf area density estimates (see Fig. 5). As mentioned above, the use of different voxel sizes yields different leaf area estimates over a given volume, and guidelines are needed for selecting an appropriate value. The LiDAR measurements of blue oak trees from the California site and of shea trees from the Malian site (the blue oak leaves being significantly smaller than the shea tree leaves) provide new insights on the choice of voxel size. To investigate the effect of voxel size on both datasets, we turn to the work of Lang and Xiang (1986) who averaged transmission observations over a horizontal path. In their study, the leaf area index of vegetation with large gaps was derived from gap fraction estimates obtained from a series of light transmission observations at fixed length intervals along a transect. Two important assumptions underlie the work of Lang and Xiang (1986), namely, (1) that foliage is randomly distributed within the volume of interest (i.e., there is no foliage clumping), and (2) that the Poisson law applies (i.e., the scatterers dimensions are much smaller than those of the volume of interest). This is also the case for retrieval of leaf area distribution from terrestrial LiDAR point clouds.

The work of Lang and Xiang (1986) is based on the theory developed by Monsi and Saeki (1953, 2005) who model light transmission through a leaf canopy by first considering the case of N flat leaves, each having a surface A_L, placed on a flat horizontal surface of area A_S. In this simplified model, leaf angles are ignored, the light trajectory is normal to the surface, and the leaves can overlap, such that the positive binomial can be used to express the probability of light going through a gap without hitting a leaf:

$$P_0 = \left(1 - \frac{A_L}{A_S}\right)^N \tag{1}$$

In this expression, $1 - A_L/A_S$ is the probability of light going through a gap when there is only one leaf on the surface. In accordance with the use of layers by Nilson (1971), N can also be interpreted as a number of layers above the horizontal surface each containing one leaf, and the positive binomial expresses the possibility of only zero or one contact per layer.

The leaf area index can then be defined as

$$\mathrm{LAI} = \frac{N \cdot A_L}{A_S} \tag{2}$$

This implies that for a fixed LAI and A_L, A_S increases with N. When N approaches infinity, Eq. (1) is equivalent to the Poisson law:

$$P_0 = e^{-\mathrm{LAI}} \tag{3}$$

The correspondence between the binomial expression and the Poisson law is proportional to the ratio $R = A_S/A_L$. This can be demonstrated by combining Eqs. (1)–(3):

$$\mathrm{LAI}_{\mathrm{Poisson}} = -\ln\left(1 - \frac{1}{R}\right)^{\mathrm{LAI}_{\mathrm{Binomial}} \cdot R} \tag{4}$$

Which is equivalent to:

$$\frac{\mathrm{LAI}_{\mathrm{Poisson}}}{\mathrm{LAI}_{\mathrm{Binomial}}} = -R \cdot \ln\left(1 - \frac{1}{R}\right) \tag{5}$$

For large values of R, the surface area is much greater than the leaf area and Eq. (5) approaches unity, meaning that the LAI estimates obtained from the Poisson and binomial expressions are equivalent. Similar to sun beam transmission observations along a transect, the LiDAR laser pulses going through a voxel can be considered as a set of probes inserted irregularly along parallel transects of length equal to the voxel side length (Δh). Mann et al. (1977) showed that in the case of a transect crossing the surface A_S, R can also be expressed as the ratio of transect length (Δh) to the mean secant of a leaf (s):

$$R = \frac{\Delta h}{s} \tag{6}$$

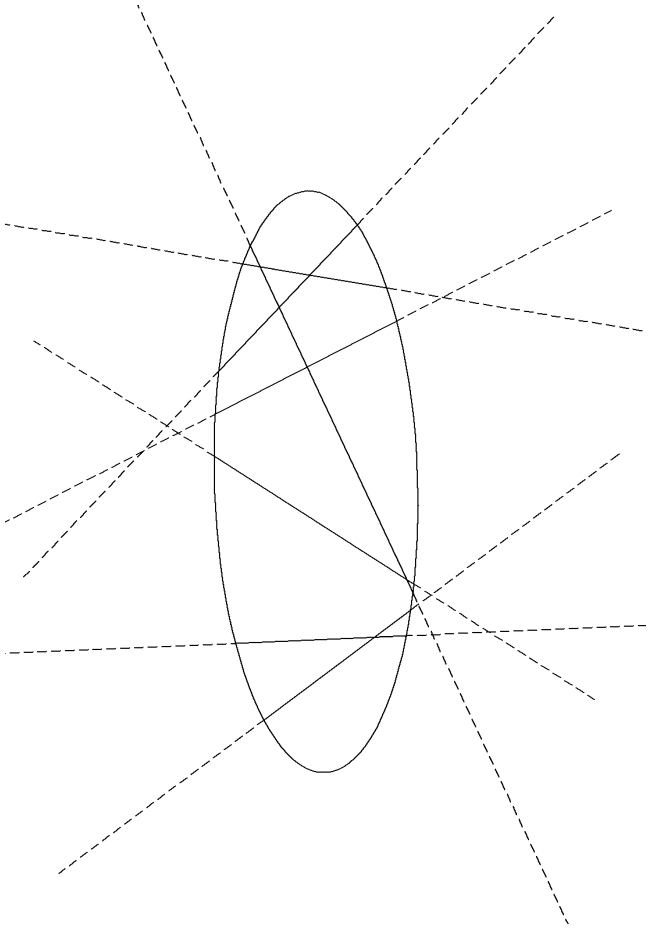

Fig. 6. Representation of the mean secant length (s) through a leaf, which corresponds to the mean length of random lines going through the plane of the leaf. In this figure, the mean secant corresponds to the mean length of all solid lines. Only a few lines are used here to illustrate the concept; to compute s a much higher number of random lines should be used.

In the above equation, s can be interpreted as the mean length of secants through the leaf, defined by the intersection of a linear transect with the leaf edges. Fig. 6 shows an example of some secants drawn through a leaf represented by an ellipse. In the figure, the mean secant length is the mean length of all continuous portions of the random transects running through the leaf. To derive s values for the shea and oak tree leaves, we assumed elliptical shapes for the leaves, and a computer program was written to compute to average secant of a series of projected ellipses. To account for leaf inclination angles in 3-D space, one needs to compute the mean secant length (s) using a leaf that is projected perpendicularly to the incidence angle of the probe. Since tree leaves have various orientations and inclinations, angle distribution functions are commonly used to represent the variations in vertical foliage angle and azimuthal orientation. These functions are generally referred to as G functions, and represent the mean projection of a unit foliage area in a particular direction of interest (Ross, 1981). Here, the direction of interest was defined by the mean elevation angle between the LiDAR instrument and the tree crowns for the oak and the shea. The leaf inclination angle distribution functions used were those best fitting leaf angle measurements at each site: spherical distribution for the shea trees, and erectophile distribution for the oak trees.

The choice of leaf angle distribution function was based on terrestrial LiDAR measurements for the shea trees at the Mali site (Béland et al., 2011), and from digital photography for the oak trees at the Tonzi site (Ryu et al., 2010). Azimuthal symmetry was assumed in the computations. The computed s values were 4.38 cm and 1.34 cm for the shea and oak tree leaves respectively.

Lang and Xiang found that a value for the transect length of 10 times the mean secant of a leaf was appropriate in different crops: large enough so the Poisson law is an acceptable approximation of the binomial expression, and small enough not to include large spaces between rows. Here we hypothesize that the same value of $R = 10$ can be used when measuring natural tree crowns, and that this can be used to determine an appropriate voxel size for sampling the transmission of the LiDAR laser pulses.

3. Results and discussion

3.1. Correcting for occlusion

The LiDAR measurements from five individual shea trees and the 20 m × 20 m oak trees plot area were analyzed with the method described in Béland et al. (2011) using voxel sizes between 5 and 200 cm. Occluded voxels were identified on the basis of having less than 15% of their volume explored by the laser pulses. Each occluded voxel was assigned a leaf area density value using two different approaches for the shea and oak trees. For the individual shea tree crowns, a leaf area density value corresponding to the mean leaf area density of all non-occluded voxels was assigned. This approach, much simpler than the one used in Béland et al. (2011), was selected to avoid introducing artifacts caused by the occlusion correction process being influenced by the voxel size. For the oak trees plot area, a strong relationship between height above the ground and voxel leaf area density was observed for the non-occluded voxels, and this relationship was used to compute the leaf area density value to assign to occluded voxels. This relationship can be explained by the trees putting out more leaves in the upper layers of this tree cluster where more light is available than in the lower layers. The relationship is shown in Fig. 7, and can be interpreted as follows: for a given voxel volume containing leaves, the average density of leaves increases with height following the exponential expression:

$$\mathrm{LAD}(h) = 0.651 \cdot e^{0.115 \cdot h} \quad (7)$$

A leaf area density value was assigned to the occluded voxels using Eq. (7) on the basis of the voxel's height above ground.

3.2. Distinguishing noise, wood, and foliage from terrestrial LiDAR measurements

On average, 81% (S.D.: 10%) of laser pulses generating multiple returns in the leaf-on scans corresponded to partial hits on branches and twigs. These returns cannot be differentiated from leaves on the sole basis of their apparent reflectance. By considering these pulses as noise, the amount of wood not differentiable from leaves is decreased by 59% (±12%). Nevertheless, this indicates that the information provided by multiple discrete returns terrestrial LiDAR instruments is useful for the purpose of improving the identification of the nature of targets. However, about 19% of these multiple returns pulses are valid contacts with leaves (number of laser returns marked as "leaves from multiple" in Fig. 1), and thus the threshold identification process needs to account for those discarded valid returns.

The computation of an optimal apparent reflectance threshold for separating leaf from wood pulse returns within the Riegl VZ-400 measurements of the ten blue oaks was based on balancing three

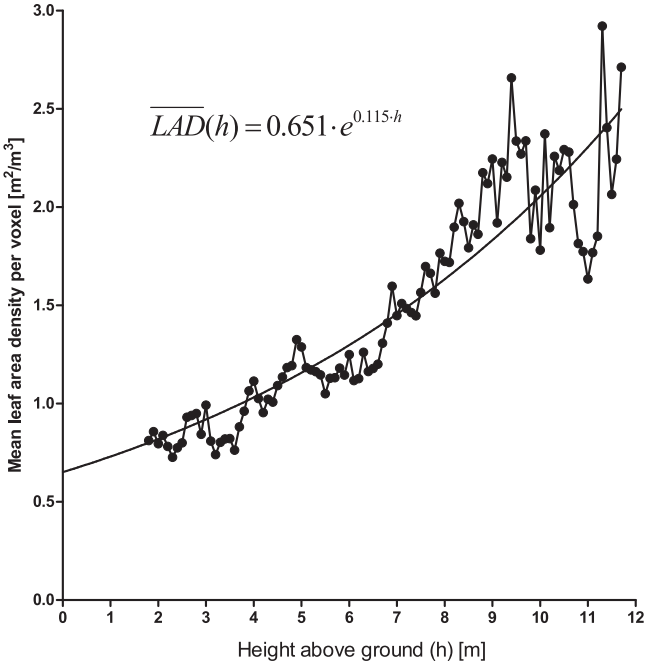

Fig. 7. Relationship between height above the ground and the mean estimates of leaf area density at the voxel level for the oak trees plot area (20 m × 20 m in size). This graph indicates that the concentration of leaves per unit volume increases with distance from the ground. The voxel size used in this example was 10 cm.

areas represented in Fig. 8: (1) the number of valid leaf returns discarded from using only pulse returns making contact with a single target (area A), (2) wood returns classified as leaves (area B + area A), and (3) leaf returns classified as wood (area C). The optimal apparent reflectance threshold is the apparent reflectance value for which the difference between the B and C areas is smallest. The lower limit of apparent reflectance corresponds to 15%, which is half of the leaf reflectance measured in the field with a spectrometer (30%) and is fixed to avoid the inclusion of pulse returns which were not centered on the leaf. The average optimal threshold for the ten trees was 28% (S.D.: 1.6%), which is slightly lower than the average measured leaf reflectance.

Identifying the proper apparent reflectance threshold requires an understanding of the interactions between laser pulses – whose characteristics are specific to each LiDAR instrument – and the plant parts within the canopy. This process benefits from having access to scans in leaf-on and leaf-off conditions, ideally from the same positions and using the same angular resolution. Having access to those for evergreen tree species may not be possible, and for that reason the recent development of dual wavelength terrestrial LiDAR instruments is of particular interest to the identification of foliage and wood within the data. The interactions between the pulses and plant parts depend on leaf size, branching architecture, leaves and bark spectral properties, and LAI/PAI ratios, which are highly site specific. They also depend on the incident angle of the laser pulses; when walking below a tree crown one can notice that the wood structure is fairly visible, and when flying at very low altitude above a forest canopy in leaf-on condition one can observe how foliage masks the wood structure and how little of it can be seen. This highlights the inherent differences in observations of a canopy from above with an airborne LiDAR, from below with a terrestrial LiDAR in a tall canopy, and from the side with a terrestrial LiDAR placed next to a relatively small tree.

Methods using information on multiple returns could be developed further by considering the intensity of reflected light from target hits occurring after a first contact; since more energy

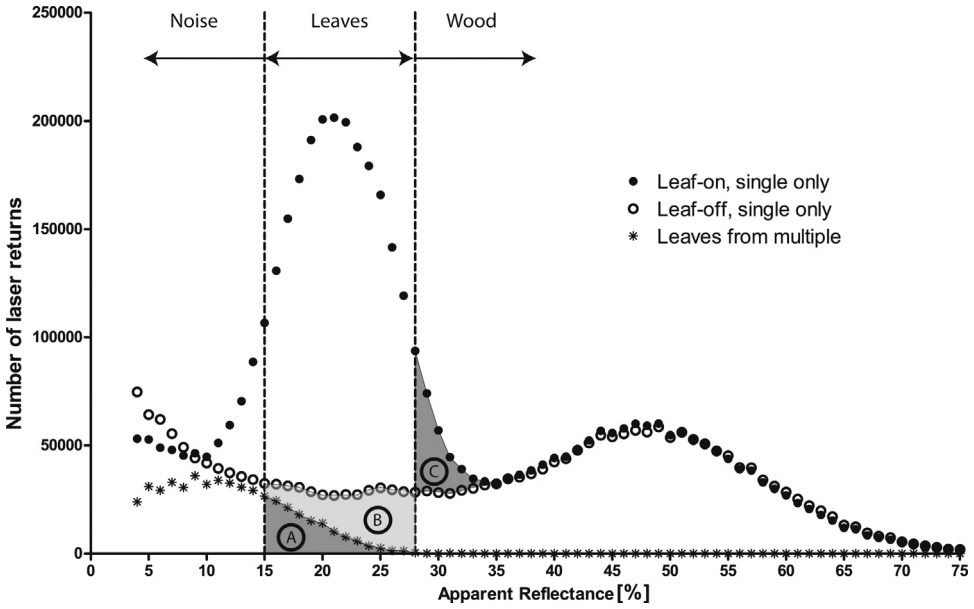

Fig. 8. Histograms of the TLS point clouds in leaf-on and leaf-off conditions including only those pulses that made contact with one single target. The method for identifying the optimal apparent reflectance threshold value for separating wood from foliage returns (vertical dashed line on the right side) considers the three areas in gray: (1) the number of leaf returns from pulses making contact with multiple targets (area A), (2) wood returns classified as leaves (area B + area A), and (3) leaf returns classified as wood (area C). The upper threshold is placed at a reflectance value where area B is nearest to area C. The lower threshold refers to half of the leaf reflectance (15% reflectance), and is used to identify partial laser pulse hits on leaves for which the pulse center falls outside the leaf area.

reflected from targets generating succeeding hits increases the likelihood of the first hit being generated from a small portion of the pulse hitting a bright target. However, because the laser pulses may carry all or part of their energy well beyond the instrument's maximum measurement range, a definite solution to the problem of pulses partially hitting targets cannot be obtained. Because LiDAR systems measure reflected light, there is no way of knowing which fraction of the pulse energy travels beyond the instrument maximum measurement range (unless the equivalent of pulse energy reflected from a fully illuminated 100% Lambertian reflectance target is measured, in which case one may assume that none or little of the pulse's energy keeps moving forward). Since the target reflectance and the portion of energy hitting the target both contribute to the observed return intensity, the true reflectance of partially hit individual targets cannot be unmixed.

Using half of the leaf reflectance as a threshold for excluding laser pulse returns not centered on leaves is expected to yield satisfactory results for leaves not having excessive anisotropic properties and having regular shapes with little indentation. More complex leaf shapes and deep indented margins may require a more thorough analysis such as the one used in Béland et al. (2011). One should also be aware of the fact that there are many (tropical) species of leaves with more than one color, e.g. white + green, or purple + green, which may complicate the analysis since the reflectance of these leaves at the laser wavelength may not be homogeneous across the leaf surface.

3.3. Voxel size

Fig. 9 shows the drip line LAI estimates (the drip line is an area defined by the outermost extent of a tree crown projected on the ground) for three of the individual shea tree crowns and for voxel sizes between 5 and 200 cm – the remaining crowns are not shown for clarity reasons; they had similar behavior. The data represents LAI values obtained using the method of Béland et al. (2011) for different voxel sizes. The LAI estimates correspond to estimates from the Poisson law, since these values were computed using the contact frequency method of Warren Wilson (1960) which makes the same assumptions as the Poisson law. Eq. (4) was fitted to the data at a value of $\Delta h/s$ corresponding to a voxel size where the effects of large gaps between branches and crowns could be assumed minimal (solid line curves in the figures). These solid line curves represent the expected deviation between the Poisson and binomial estimates. The solid lines converge toward values which are indicated using horizontal dotted lines, corresponding to the LAI value where both the Poisson and binomial estimates are expected to agree. Fig. 10 shows the plot level LAI estimates for the 20 m × 20 m oak trees plot and for the same voxel size range.

For the shea trees (Fig. 9), the data for voxel sizes 10, 20 and 30 cm fit well the theoretical curve from Eq. (4) for the low LAI tree (Tree 1), but as the LAI increases the data for the 10 and 30 cm sizes deviates from the theoretical curve (solid lines versus dashed lines). In the case of the oak trees, the data for voxel size above 10 cm deviate more rapidly from the theoretical curve. This is because voxel sizes of 20 cm and above contain larger gaps for the oak trees than the shea trees.

It can be seen from Figs. 9 and 10 that the LAI estimates are lower for very small voxel sizes (below 10 cm), reach a maximum for 10 cm voxels, and decrease again as voxel size increases. For shea Trees 1 and 3, the LAI values which correspond to the estimate where the binomial expression and the Poisson law converge (i.e. where Eq. (5) equals unity) are slightly lower than the LAI values obtained from direct measurements involving leaf harvesting (for further details on those measurements see Béland et al., 2011), and for shea Tree 5, the difference is more significant. This may be explained by the simple treatment of occlusion effects done here, and emphasizes the importance of occlusion effects in densely foliaged crowns. Again, a simple approach to correcting for occlusion effects was adopted here in order to isolate the effect of voxel size on the leaf area estimates and minimize the impact of a particular occlusion correction procedure. For the blue oak trees, the LAI value which corresponds to the estimate where the binomial expression and the Poisson law converge (0.71) is slightly lower that the peak landscape level LAI value estimated from the litterfall method of 0.82 reported in Ryu et al. (2010) for the Tonzi site. This difference may be linked to two factors: (1) the area

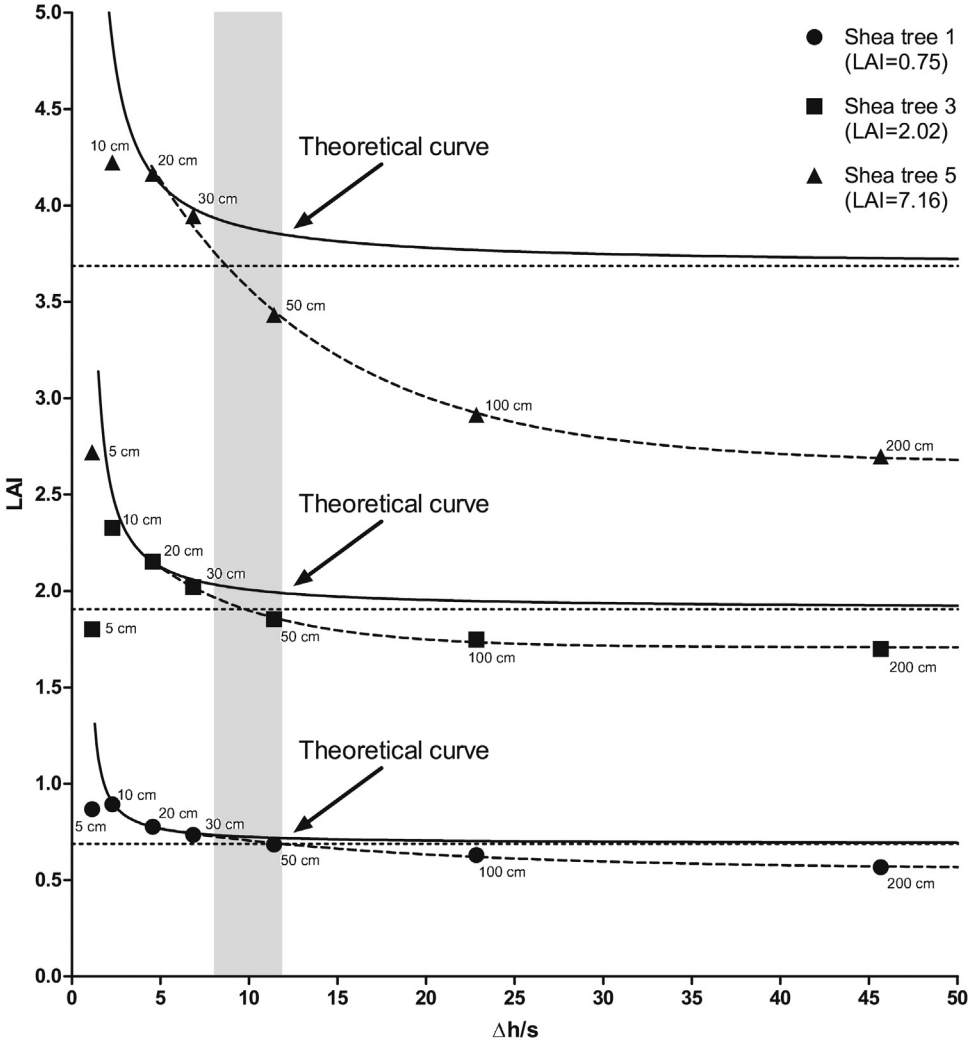

Fig. 9. Crown level leaf area index (drip line LAI) estimates for three of the shea trees as a function of voxel size (Δh). The value of s for the shea trees is 4.38 cm. The values in cm marked refer to the voxel sizes. The solid curves marked 'theoretical curve' refer to Eq. (4) fitted to the 20 cm voxel size data, where $R = \Delta h/s$. The horizontal dotted lines refer to the value of $LAI_{Binomial}$ in Eq. (4), i.e. the estimated correct LAI value since it refers to the LAI estimate where the binomial expression and the Poisson law converge. The drip line LAI values given in the legend for each tree are those obtained from direct measurements (leaf harvesting) and published in Béland et al. (2011). The shaded area indicates range of values for Δh/s between 8 and 12. The dashed curves are used solely to support the interpretation of the figure.

sampled was different, as the litter was collected at 25 locations spread over the landscape and the LiDAR-based LAI estimate here focuses on a 20 m × 20 m area, and (2) the litter was collected in 2008 and the LiDAR data was acquired in 2012, and it is possible the drier than average winter conditions in 2011–2012 may have affected the amount of leaves produced that year relative to 2008.

The ability to detect occlusion effects remains an important factor if one wishes to map the 3-D distribution of leaf area. When identifying occluded voxel volumes, it is important to distinguish between occluded voxels containing material and occluded voxels, which are empty and correspond to gaps between crowns or between branches and do not contain leaves. Our approach poses two conditions in the identification of an occluded voxel volume: (1) less than 15% of the voxel volume was explored, and (2) at least one pulse made contact with a leaf. If no material is detected inside a voxel, it is ignored when correcting for occlusion effects. An efficient occlusion detection process should thus be able not only to locate areas sampled by too few laser pulses, but also distinguish between occluded areas that contain material from those which do not. For this reason, the choice of voxel size should consider occlusion detection within the LiDAR data.

The decrease in LAI estimates at smaller voxel sizes contradicts the behavior of Eq. (4), which predicts increasing values of LAI as voxel size gets smaller. As voxel size gets smaller, fewer pulses can enter the voxel, causing the statistics on transmission to be less reliable. Cases where transmission is null are more frequent. For such cases, leaf area density estimates cannot be obtained. A reduced number of pulses entering the voxel also impede the localization of occluded areas. For example, if the lateral distance between consecutive pulses is 1 cm and the voxel size is 5 cm, the maximum number of pulses capable of entering the voxel from one of its sides is 25, and assuming less than 15% of those laser pulse entered the voxel would result in fewer than 4 pulses actually entering the voxel. Most terrestrial LiDAR instruments are capable of making measurements with distances between laser pulses slightly smaller than 1 cm from a typical range between the instrument and the trees, but the limited amount of pulses available to compute statistics at the voxel scale and the cross-sectional size of the laser pulse are deemed limiting factors in using very small voxels for describing the 3-D spatial distribution of leaf area.

At larger voxel sizes, two factors – somewhat related – combine to yield smaller LAI estimates. The first relates again to a difficulty in detecting occlusion effects, because as voxels get larger, the

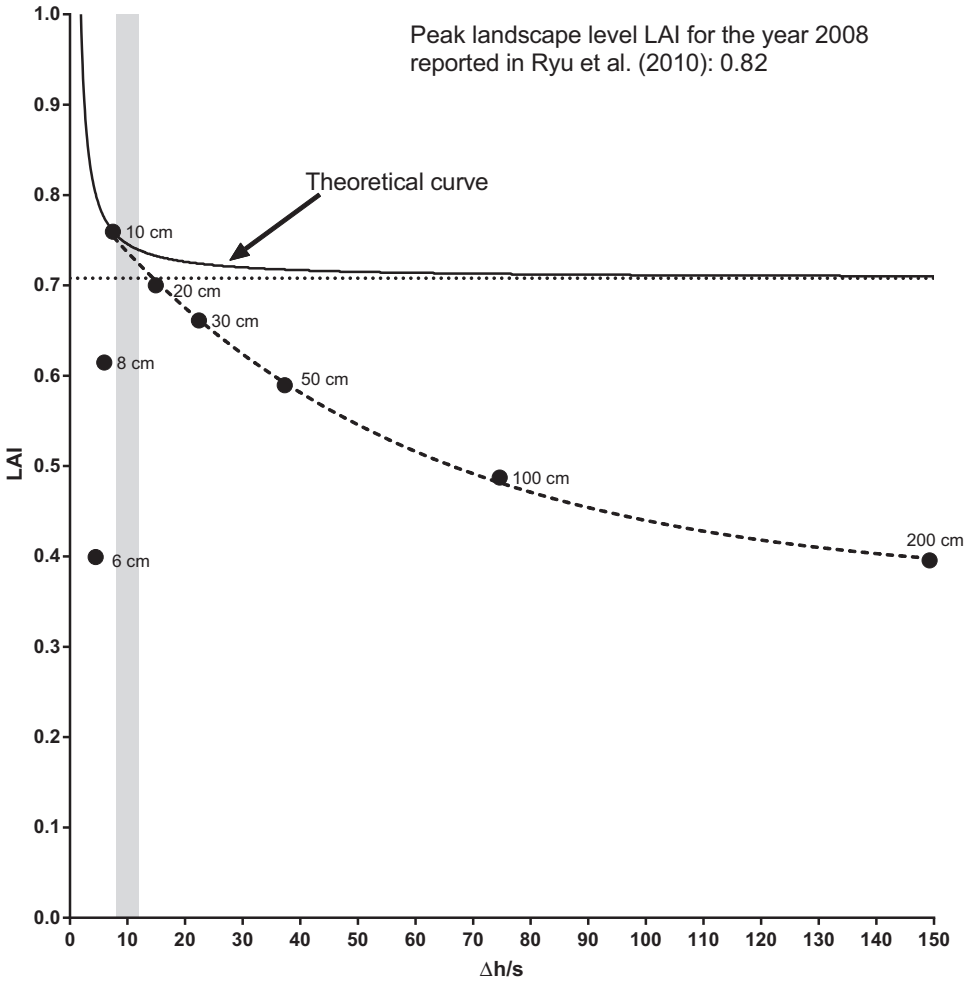

Fig. 10. Plot level leaf area index (LAI) estimates for the oak Trees 20 m × 20 m plot area for different voxel sizes (Δh). The value of s for the shea trees is 1.34 cm. The solid curves marked 'theoretical curve' refer to Eq. (4) fitted to the 10 cm voxel size data, where $R = \Delta h/s$. The horizontal dotted lines refer to the value of $LAI_{Binomial}$ in Eq. (4), i.e. the estimated correct LAI value since it refers to the LAI estimate where the binomial expression and the Poisson law converge. The shaded area indicates range of values for $\Delta h/s$ between 8 and 12.

probability of less than 15% of its entire volume being explored decreases as gaps in the canopy smaller than the voxel will allow most pulses to keep moving forward. More importantly, as voxel size increases, so does the probability of the volume including large gaps, and these gaps invalidate the assumption of random distribution within the voxel. Chen and Black (1992) state that regardless of how small the sampling volume is, the effect of foliage clumping is inevitable and suggest that a correction factor for a given volume size be determined. Ongoing experimental research on the blue oak trees indicates that the distribution of foliage at the 30 cm voxel scale varies from clumped to regular, but that on average the distribution can be considered random.

To investigate the detection of occluded volumes, the fraction of occluded voxels was computed for the different voxel sizes for the shea trees and the oak trees plot area. For this analysis we considered only the voxels within which at least one contact between a laser pulse and a leaf was recorded – indicating the presence of material inside the voxel. The occluded voxels refer to those voxels where less than 15% of the volume was explored. The results, shown in Fig. 11, indicate that the maximum detected occlusion fractions correspond to voxel sizes between about 5 and 20 cm. These maxima suggest that the fraction of occluded voxels increases with three factors: (1) the foliage density of single crowns, (2) the size of the scanned area, and (3) the number of scan locations from where the LiDAR measurements were made and their spatial resolution (angle between consecutive laser pulses). The decrease in the detected occluded fraction observed for small voxels is caused by the low number of pulses entering the voxels, this is less of an issue where occlusion effects are low (e.g. shea Tree 1). Larger voxel sizes are unable to pinpoint the location of the occluded areas, and are unable to distinguish between occluded voxel with and without plant material inside. It is important to note that the optimal voxel size for locating occlusion effects may not be directly related to leaf size but rather to the distance between consecutive laser pulses and the predominance of occlusion effects. In this study the LiDAR measurements were made using a similar angle between consecutive measurements (260–280 μrad) which resulted in a distance between pulses of about 5 mm at a distance of 20 m from the instrument. Using this resolution, our results indicate that voxel sizes between 5 and 20 cm are preferable for locating occlusion effects in a wide range of foliage densities and measurement conditions, and that the optimal voxel size is dependent on the predominance of occlusion effects, as high levels of occlusion are associated with reduced numbers of unintercepted pulses, which in turn prevent the use of smaller size voxels.

Because occlusion effects can significantly alter the leaf area estimates, a subsample of the 20 m × 20 m × 12 m oak tree plot area where occlusions were very low was selected to compute leaf area estimates. The selected area is 4 m × 4 m × 4 m in size. It was located

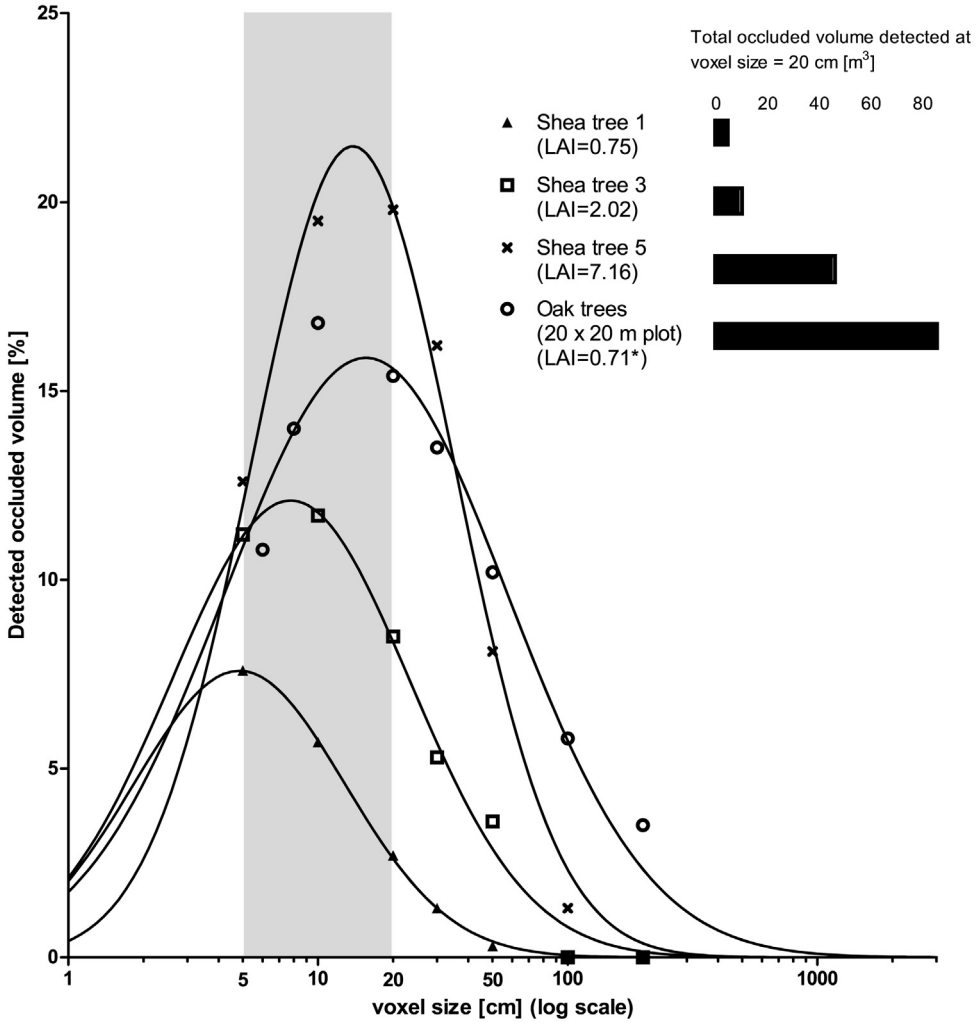

Fig. 11. Percentage of voxels where leaf material was detected which are considered occluded as a function of voxel size. Note the log scale on the abscissa. A voxel is considered occluded when less than 15% of its volume was explored by the laser pulses. An indication of the total volume occluded for each of the shea tree and for the oak trees area is given in the legend. Also in the legend are the LAI estimates obtained from direct measurement for the shea trees, and the LAI estimate obtained from Fig. 6 for the oak trees plot area (marked with an *). The graph indicates that as occlusion effects increase the optimal voxel size for detecting the occluded also increases. This is explained by the reduced number of pulses available to enter smaller voxels when occlusion effects are large.

next to a clearing within the larger plot, allowing a majority of the laser pulses emitted from three of the six scanning positions to reach it. The rationale for using this subset volume was to isolate the effect of large gaps on the LAI estimates. Fig. 12 shows the LAI estimates for the smaller area for different voxel sizes. The data for small voxel sizes closely fits the theoretical curve for Eq. (4) since occlusion is negligible. The small gap between the data points and the curve at small voxel sizes may be due to inaccuracies in the estimation of the s parameter computed for the oak trees. The figure indicates that for voxel sizes of 20 and 30 cm, the presence of large gaps inside the voxels results in an underestimation of about 10–15% in the LAI estimate for the oak trees. By contrast, the lowest LAI shea tree data for the same voxel sizes closely fits the curve for Eq. (4). For both the shea and oak trees, the curve fitted to the data for the different voxel sizes crosses the horizontal dotted line (which refers to the LAI estimate from the binomial expression) at values of $\Delta h/s$ slightly below 10.

For the subsample oak tree volume in Fig. 12 and the shea Tree 1 in Fig. 9 – both being little affected by occlusion effects – the LAI estimates match the theoretical curve of Eq. (4) for voxel sizes below 10 times the leaf size ($\Delta h/s < 10$). For $\Delta h/s$ values above 10, the LAI estimates deviate from the theoretical curve because spaces between branches and crowns are being included with leaves within the sampled voxel volumes, causing an underestimation of LAI. This is expected to mainly result from clumping within the larger voxel volume as mentioned above.

The optimal voxel size should thus allow (1) detecting occluded voxels, (2) getting a reasonable match between the Poisson and binomial expressions, and (3) exclude large spaces between branches and crowns. Our results indicate that a voxel size slightly below 10 times the leaf size (or mean projected secant through the leaf) meets requirements (2) and (3). Requirement (1) is dependent on the LiDAR scanning configuration and the canopy environment. The choice of a suitable voxel size for estimating the 3-D spatial distribution of leaf area in broadleaf forest canopies should consider all 3 requirements above. For the oak and shea trees used in this study, our results indicate that in order to meet requirement (1) voxel sizes below 20 cm should ideally be used. For the oak trees, using voxels of 10 cm would satisfy requirement (3), and by multiplying the LAI estimates by 0.93 would correct for the divergence from the binomial expression from using the Poisson law (10 cm voxels for the oak trees corresponds to $R = 7.5$, for which Eq. (5) yields 1.07, meaning that the estimate from the Poisson law is overestimated by 7%). For the shea trees, using voxels 30 cm in size and multiplying the LAI estimates by a factor of 0.93 ($R = 6.85$ and Eq. (5) yields 1.08) appears to be the best compromise toward

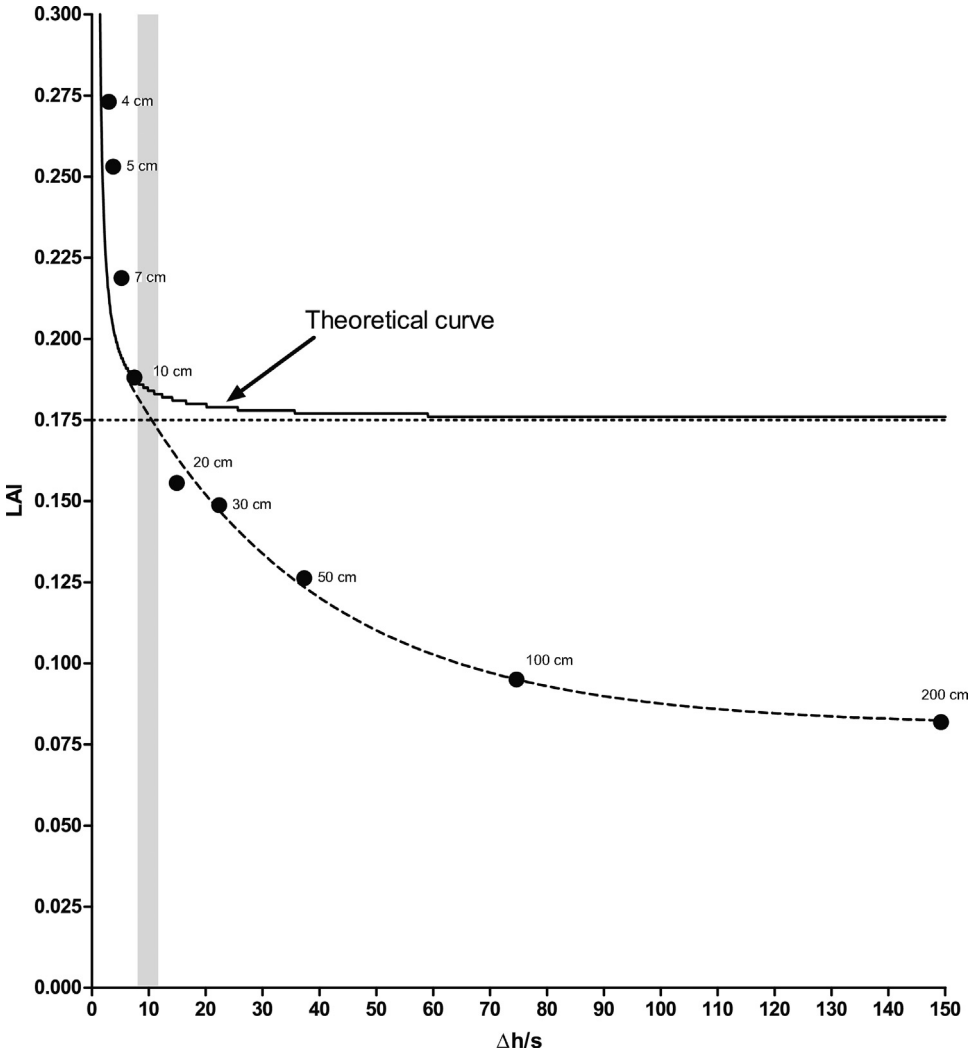

Fig. 12. Sub-plot level leaf area index estimates over the 4 m × 4 m × 4 m area (where occlusion effects were negligible) for the oak trees as a function of voxel size (Δh). The value of s for the shea trees is 1.34 cm. The solid curves marked 'theoretical curve' refer to Eq. (4) fitted to the 10 cm voxel size data, where $R = \Delta h/s$. The horizontal dotted lines refer to the value of $LAI_{Binomial}$ in Eq. (4), i.e. the estimated correct LAI value since it refers to the LAI estimate where the binomial expression and the Poisson law converge. The shaded area indicates range of values for $\Delta h/s$ between 8 and 12.

meeting all three requirements. In all cases, as occlusion effects increase so does the importance of using a voxel size appropriate for detecting the occluded areas. Conversely, when occlusion effects are negligible (e.g., shea Tree 1) considerations for the choice of voxel size should essentially relate to leaf size. The results obtained here show the importance of minimizing occlusion effects through appropriate field measurement protocols in order to allow the use of a voxel size best adapted to the size of the leaves in the canopy.

4. Conclusion

Detailed information on the spatial distribution of leaf area in forested environments can provide an exceptional basis for studies on ecosystem and agro-ecosystem functioning, as well as investigations to validate satellite remote sensing procedures to retrieve important surface parameters like leaf area density. In this study, we first tested the suitability of multiple discrete return terrestrial LiDAR measurements for improving the identification of laser pulse returns originating from wood and leaves. Distinguishing between pulse returns from wood and leaves is essential to directly derive leaf area density instead of plant area density. The method used here involved the use of tree scans in leaf-off and leaf-on conditions, which proved to be crucial for such an analysis. The results indicated that information from multiple discrete return LiDAR, which became only available with the latest generation of instruments, holds potential for supporting the interpretation of the reflected energy's intensity.

We estimated drip line leaf area index (LAI) values at the individual crown and small plot level using measurements from two savanna sites, one in Mali and the other in California, that were populated with different tree species having contrasting leaf sizes. This study permitted a better understanding of the implications of voxel size on leaf area estimates, as optimal voxel size is in part linked to leaf size. The optimal voxel size should allow detecting occluded voxels, getting a reasonable match between the Poisson and binomial expressions, and excluding large spaces between branches and crowns. We found that voxel linear dimensions of about 10 times the leaf size are suitable for minimizing the presence of large gaps within the voxel, which reduces the risks of violating the assumption of random distribution. Using such voxel size can result in an overestimation of the leaf area estimates because the Poisson law deviates from the positive binomial distribution. This overestimation is in the order of 5% and can be corrected for by multiplying the leaf area estimates by a factor of 0.95. It is important to note that the overestimation is not related to the distribution of foliage

within the voxel (i.e. clumped or regular), but strictly to the size of the leaves with regards to the size of the sampled volume. If occlusion effects are significant, the use of a smaller voxel size may be required to pinpoint the location of the occluded areas. In such case, the correction factor of 0.95 should be adjusted to account for the larger deviation between the Poisson law and the binomial expression. It is however preferable to minimize occlusion effects through appropriate field measurements protocols and allow the use of a voxel size better adapted to the size of the leaves and the branching structure of the canopy.

It is known that occlusion is a critical issue in the estimation of leaf area density from LiDAR measurements, and the results presented here show that the selected voxel size influences the capacity to pinpoint the location of the occluded volumes. Few studies have investigated methods for minimizing occlusion effects in terrestrial LiDAR measurements of forest canopies. Accounting for occlusion effects thus remains a major challenge and particular attention should be given to field procedures during the terrestrial LiDAR measurements which determine the level of occlusion. Occlusion effects also need to be better understood and addressed in data processing methods. Further research is thus needed to improve methods for correcting for these effects, as well as generate guidelines for field measurement protocols aiming at minimizing occlusions in different canopy environments.

Using terrestrial LiDAR measurements to derive spatially explicit leaf area estimates opens up new opportunities which merit more scrutiny and methodological development. The results obtained with field instruments could help calibrate and validate products derived from similar airborne or space borne sensors, as well as support investigations of the exchanges of water and other biogeochemicals between plants and the atmosphere. For such LiDAR-based technique to be truly useful it is critical to get the leaf area estimates correct at the voxel scale. Our findings provide some guiding principles on the choice of voxel size. We suggest that studies aiming at improving these techniques give due consideration to the characteristics of the LiDAR instrument used, in particular to the laser footprint size which has great influence on the estimation of gap fraction in forest canopies.

Acknowledgments

This research was supported by the *Fond Québécois de la Recherche sur la Nature et les Technologies* (FQRNT), the US Department of Energy Terrestrial Carbon Program, grant No. DE-SC0005130, the Natural Sciences and Engineering Research Council of Canada (NSERC) (through the Collaborative Research and Development (CRD) grant program in partnership with Boreal-Information Strategies), and the Global Environmental and Climate Change Centre (GEC3) International Internship Program. The authors thank Jean-François Côté and Simon Chamberland for their work on the computation of voxel statistics from terrestrial LiDAR point clouds using the LVox model. We are also grateful to Marcelin Sanou, Baga Samake, and the Mali directorate of the *Service des Eaux et Forêts* for their support with field activities on the Mali site, and to Craig Glennie and Darren Hauser from the University of Houston and the NCALM center for their assistance with terrestrial LiDAR measurements at the Tonzi site. The authors acknowledge the support of their respective home institutions for this interdisciplinary and international collaboration, and MB thanks Alan Belward and the Institute for Environment and Sustainability of the Joint Research Centre for hosting him in 2010 and 2011.

References

Asner, G.P., Hughes, R.F., Vitousek, P.M., Knapp, D.E., Kennedy-Bowdoin, T., Boardman, J., Martin, R.E., Eastwood, M., Green, R.O., 2008. Invasive plants transform the three-dimensional structure of rain forests. Proc. Natl. Acad. Sci. U. S. A. 105 (11), 4519–4523.

Asner, G.P., Scurlock, J.M.O., Hicke, J.A., 2003. Global synthesis of leaf area index observations: implications for ecological and remote sensing studies. Global Ecol. Biogeogr. 12 (3), 191–205.

Baldocchi, D.D., Harley, P.C., 1995. Scaling carbon-dioxide and water-vapor exchange from leaf to canopy in a deciduous forest. 2. Model testing and application. Plant Cell Environ. 18 (10), 1157–1173.

Balduzzi, M.A.F., Van der Zande, D., Stuckens, J., Verstraeten, W.W., Coppin, P., 2011. The properties of terrestrial laser system intensity for measuring leaf geometries: a case study with conference pear trees (*Pyrus communis*). Sensors 11 (2), 1657–1681.

Béland, M., Widlowski, J.-L., Fournier, R., Côté, J.-F., Verstraete, M.M., 2011. Estimating leaf area distribution in savanna trees from terrestrial LiDAR measurements. Agric. For. Meteorol. 151 (9), 1252–1266.

Cannell, M.G.R., 1989. Physiological basis of wood production: a review. Scand. J. For. Res. 4, 459–490.

Cescatti, A., 1997. Modelling the radiative transfer in discontinuous canopies of asymmetric crowns. I. Model structure and algorithms. Ecol. Model. 101 (2), 263–274.

Chen, J.M., Black, T.A., 1991. Measuring leaf-area index of plant canopies with branch architecture. Agric. For. Meteorol. 57 (1–3), 1–12.

Chen, J.M., Black, T.A., 1992. Foliage area and architecture of plant canopies from sunfleck size distributions. Agric. For. Meteorol. 60 (3–4), 249–266.

Chen, Q., Baldocchi, D., Gong, P., Kelly, M., 2006. Isolating individual trees in a savanna woodland using small footprint LiDAR data. Photogramm. Eng. Remote Sens. 72 (8), 923–932.

Cohen, S., Fuchs, M., 1987. The distribution of leaf-area, radiation, photosynthesis and transpiration in a shamouti orange hedgerow orchard. 1. Leaf-area and radiation. Agric. For. Meteorol. 40 (2), 123–144.

Danson, F.M., Hetherington, D., Morsdorf, F., Koetz, B., Allgower, B., 2007. Forest canopy gap fraction from terrestrial laser scanning. IEEE Geosci. Remote Sens. Lett. 4 (1), 157–160.

Fukai, S., Loomis, R.S., 1976. Leaf display and light environments in row-planted cotton communities. Agric. Meteorol. 17 (5), 353–379.

Gastellu-Etchegorry, J.P., Demarez, V., Pinel, V., Zagolski, F., 1996. Modeling radiative transfer in heterogeneous 3-D vegetation canopies. Remote Sens. Environ. 58 (2), 131–156.

Hosoi, F., Nakai, Y., Omasa, K., 2010. Estimation and error analysis of woody canopy leaf area density profiles using 3-D airborne and ground-based scanning LiDAR remote-sensing techniques. IEEE Trans. Geosci. Remote Sens. 48 (5), 2215–2223.

Hosoi, F., Omasa, K., 2006. Voxel-based 3-D modeling of individual trees for estimating leaf area density using high-resolution portable scanning LiDAR. IEEE Trans. Geosci. Remote Sens. 44 (12), 3610–3618.

Hosoi, F., Omasa, K., 2007. Factors contributing to accuracy in the estimation of the woody canopy leaf area density profile using 3D portable LiDAR imaging. J. Exp. Bot. 58 (12), 3463–3473.

Hosoi, F., Omasa, K., 2009. Detecting seasonal change of broad-leaved woody canopy leaf area density profile using 3D portable LiDAR imaging. Funct. Plant Biol. 36, 998–1005.

Hutchison, B.A., Matt, D.R., McMillen, R.T., Gross, L.J., Tajchman, S.J., Norman, J.M., 1986. The architecture of a deciduous forest canopy in eastern Tennessee, USA. J. Ecol. 74 (3), 635–646.

Iio, A., Kakubari, Y., Mizunaga, H., 2011. A three-dimensional light transfer model based on the vertical point-quadrant method and Monte-Carlo simulation in a *Fagus crenata* forest canopy on Mount Naeba in Japan. Agric. For. Meteorol. 151 (4), 461–479.

Jackson, J.E., Palmer, J.W., 1972. Interception of light by model hedgerow orchards in relation to latitude, time of year and hedgerow configuration and orientation. J. Appl. Ecol. 9 (2), 341.

Jackson, J.E., Palmer, J.W., 1979. Simple-model of light transmission and interception by discontinuous canopies. Ann. Bot. 44 (3), 381–383.

Jacquemoud, S., Ustin, S.L., 2001. Leaf optical properties: A state of the art. In: 8th International Symposium of Physical Measurements & Signatures in Remote Sensing, pp. 223–332.

Jarvis, P., Miranda, H., Muetzelfeldt, R., 1985. Modelling canopy exchanges of water vapour and carbon dioxide in coniferous forest plantations. In: Hutchison, B.A., Hicks, B.B. (Eds.), The Forest–Atmosphere Interaction. Reidel, Dordrecht, pp. 521–542.

Jonckheere, I., Fleck, S., Nackaerts, K., Muys, B., Coppin, P., Weiss, M., Baret, F., 2004. Review of methods for in situ leaf area index determination – Part I. Theories, sensors and hemispherical photography. Agric. For. Meteorol. 121 (1–2), 19–35.

Jupp, D.L.B., Culvenor, D.S., Lovell, J.L., Newnham, G.J., Strahler, A.H., Woodcock, C.E., 2009. Estimating forest LAI profiles and structural parameters using a ground-based laser called 'Echidna (R)'. Tree Physiol. 29 (2), 171–181.

Kobayashi, H., Baldocchi, D.D., Ryu, Y., Chen, Q., Ma, S., Osuna, J.L., Ustin, S.L., 2012. Modeling energy and carbon fluxes in a heterogeneous oak woodland: a three-dimensional approach. Agric. For. Meteorol. 152 (2), 83–100.

Koike, F., 1985. Reconstruction of two-dimensional tree and forest canopy profiles using photographs. J. Appl. Ecol. 22 (3), 921–929.

Kucharik, C.J., Norman, J.M., Gower, S.T., 1998. Measurements of branch area and adjusting leaf area index indirect measurements. Agric. For. Meteorol. 91 (1–2), 69–88.

Landry, R., Fournier, R., Ahern, F., Lang, R., 1997. Tree vectorization: a methodology to characterize fine tree architecture in support of remote sensing models. Can. J. Remote Sens. 23 (2), 91–107.

Lang, A.R.G., Xiang, Y.Q., 1986. Estimation of leaf-area index from transmission of direct sunlight in discontinuous canopies. Agric. For. Meteorol. 37 (3), 229–243.
Lovell, J.L., Jupp, D.L.B., Culvenor, D.S., Coops, N.C., 2003. Using airborne and ground-based ranging LiDAR to measure canopy structure in Australian forests. Can. J. Remote Sens. 29 (5), 607–622.
Ma, S.Y., Baldocchi, D.D., Xu, L.K., Hehn, T., 2007. Inter-annual variability in carbon dioxide exchange of an oak/grass savanna and open grassland in California. Agric. For. Meteorol. 147 (3–4), 157–171.
Mann, J.E., Curry, G.L., Hartfiel, D.J., Demichele, D.W., 1977. General law for direct sunlight penetration. Math. Biosci. 34 (1–2), 63–78.
Medlyn, B., 2004. A maestro retrospective. Forests at the land–atmosphere interface. CAB Int., 105–121.
Monsi, M., Saeki, T., 1953. Über den lichtfaktor in den pflanzengesellschaften und seine bedeutung für die stoffproduktion. Jpn. J. Bot. 14, 22–52.
Monsi, M., Saeki, T., 2005. On the factor light in plant communities and its importance for matter production. Ann. Bot. 95 (3), 549–567.
Moorthy, I., Miller, J.R., Hu, B.X., Chen, J., Li, Q.M., 2008. Retrieving crown leaf area index from an individual tree using ground-based LiDAR data. Can. J. Remote Sens. 34 (3), 320–332.
Nilson, T., 1971. A theoretical analysis of the frequency of gaps in plant stands. Agric. Meteorol. 8 (C), 25–38.
Norman, J., Welles, J., 1983. Radiative transfer in an array of canopies. Agron. J. 75 (3), 481–488.
Norman, J.M., Campbell, G.S., 1989. Canopy structure. In: Pearcy, R.W., Ehleringer, J., Mooney, H.A., Rundel, P.W. (Eds.), Plant Physiological Ecology – Field Methods and Instrumentation. Chapman and Hall, London, pp. 301–325.
ORNL DAAC, 2012. Oak Ridge National Laboratory Distributed Active Archive Center. FLUXNET Web Page. ORNL DAAC, Oak Ridge, TN, USA, Available http://fluxnet.ornl.gov (accessed 04.09.12).
Parker, G.G., 1995. Structure and microclimate of forest canopies. In: ST1 (Ed.), Forest Canopies. Academic Press, California, USA, pp. 73–106.
Pearcy, R.W., Yang, W., 1998. The functional morphology of light capture and carbon gain in the Redwood forest understorey plant *Adenocaulon bicolor* Hook. Funct. Ecol. 12 (4), 543–552.
Pfennigbauer, M., Ullrich, A., 2010. Improving quality of laser scanning data acquisition through calibrated amplitude and pulse deviation measurement. In: Proceedings of the SPIE Defense, Security & Sensing, Orlando, FL, April 2010 (paper no. 7684-53).
Pretzsch, H., 2009. Forest Dynamics, Growth and Yield: From Measurement to Model. Springer, Berlin, 664 pp.
Ross, J., 1981. The Radiation Regime and Architecture of Plant Stands. Junk Publishers, The Hague, 391 pp.
Ryu, Y., Sonnentag, O., Nilson, T., Vargas, R., Kobayashi, H., Wenk, R., Baldocchi, D.D., 2010. How to quantify tree leaf area index in an open savanna ecosystem: a multi-instrument and multi-model approach. Agric. For. Meteorol. 150 (1), 63–76.
Sellers, P.J., Dickinson, R.E., Randall, D.A., Betts, A.K., Hall, F.G., Berry, J.A., Collatz, G.J., Denning, A.S., Mooney, H.A., Nobre, C.A., Sato, N., Field, C.B., HendersonSellers, A., 1997. Modeling the exchanges of energy, water, and carbon between continents and the atmosphere. Science 275 (5299), 502–509.
Sinoquet, H., Le Roux, X., Adam, B., Ameglio, T., Daudet, F., 2001. RATP: a model for simulating the spatial distribution of radiation absorption, transpiration and photosynthesis within canopies: application to an isolated tree crown. Plant Cell Environ. 24 (4), 395–406.
Sinoquet, H., Rivet, P., 1997. Measurement and visualization of the architecture of an adult tree based on a three-dimensional digitising device. Trees Struct. Funct. 11 (5), 265–270.
Sinoquet, H., Thanisawanyangkura, S., Mabrouk, H., Kasemsap, P., 1998. Characterization of the light environment in canopies using 3D digitising and image processing. Ann. Bot. 82 (2), 203–212.
Takeda, T., Oguma, H., Sano, T., Yone, Y., Fujinuma, Y., 2008. Estimating the plant area density of a Japanese larch (*Larix kaempferi* Sarg.) plantation using a ground-based laser scanner. Agric. For. Meteorol. 148 (3), 428–438.
Van der Zande, D., Mereu, S., Nadezhdina, N., Cermak, J., Muys, B., Coppin, P., Manes, F., 2009. 3D upscaling of transpiration from leaf to tree using ground-based LiDAR: application on a Mediterranean Holm oak (*Quercus ilex* L.) tree. Agric. For. Meteorol. 149, 1573–1583.
Wang, Y., Jarvis, P., 1990. Description and validation of an array model – MAESTRO. Agric. For. Meteorol. 51 (3), 257–280.
Wang, Y.P., Jarvis, P.G., Benson, M.L., 1990. 2-Dimensional needle-area density distribution within the crowns of pinus-radiata. For. Ecol. Manage. 32 (2–4), 217–237.
Warren Wilson, J., 1960. Inclined point quadrats. New Phytol. 59 (1), 1–8.
Warren Wilson, J., 1963. Errors resulting from thickness of point quadrats. Aust. J. Bot. 11, 178–188.
Weiss, M., Baret, F., Smith, G.J., Jonckheere, I., Coppin, P., 2004. Review of methods for in situ leaf area index (LAI) determination. Part II. Estimation of LAI, errors and sampling. Agric. For. Meteorol. 121 (1–2), 37–53.
Whitehead, D., Grace, J.C., Godfrey, M.J.S., 1990. Architectural distribution of foliage in individual pinus-radiata d don crowns and the effects of clumping on radiation interception. Tree Physiol. 7 (1–4), 135–155.
Widlowski, J.-L., Lavergne, T., Pinty, B., Verstraete, M., Gobron, N., 2006. Rayspread: a virtual laboratory for rapid BRF simulations over 3-D plant canopies. In: Graziani, F. (Ed.), Computational Methods in Transport. Springer, Berlin, pp. 211–231.